高可用系统的技术与应用

The Techniques and Application of High Availability System

蒋谢彬　李献球　(北京市5136信箱　100094)

摘要: 随着计算机应用的高速发展，对容错计算机的需求也日益迫切。而高可用计算机系统以其良好的性价比成为首选。本文通过一个应用实例，研究了高可用计算机系统的设计实现方法。

关键词: 高可用系统　集群　磁盘阵列　故障检测　故障恢复

1 引言

随着社会的发展和科技的进步，计算机已经进入人类活动的各个角落。特别是近年来，Internet的普及和电子商务的高速发展，对计算机的性能和可靠性的需求也日益增高。高可用计算机系统是指在不需要操作者干预的情况下，能够防止故障或从故障中恢复的计算机系统。在一些关键的应用中，对计算机系统的要求往往二十四小时连续工作，未预料的故障或停机将造成不可估计的损失。因此，计算机系统的高可用性实现技术成为一个研究热点。本文对计算机系统高可用性技术进行了深入的研究和探讨，并给出一个综合利用各种高可用性技术实现的工程实例。

2 系统设计中的高可用性技术

高可用计算机系统在系统设计时就应考虑采取措施从硬件平台上保证系统的可靠性。下文对常用的几种高可用技术进行了研究。

2.1 冗余技术

系统的可靠性，一个常用的方法就是采用冗余措施。冗余一般可以分为两个层次，系统级冗余和部件级冗余。

部件级冗余是提高可用性的基本方法，通常是对电源、硬盘、风扇和网卡等易发生故障且给系统造成危害最大的那些部件添加冗余配置，并设计如热插拔等类似的方便的更换机构，同时也要使系统能及时恢复到正常的部件冗余程度。

2.2 多机技术

系统级冗余是提高系统可靠性的有效途径。在工程实践中，常见的是同构型计算机双系统冗余。本文主要介绍以下三种双系统方案。

2.2.1 互援备份方案(Dual Active)

双机互援备份是指正常情况下两台主机均为工作机，同时对外提供服务，并互相监视对方机的健康情况。当一台主机出现故障，另一主机主动接管异常机的工作，以保证系统能够继续提供服务，达到不停机的目的，但正常运行的主机负载会有所增加。此时必须尽快将异常机修复以缩短正常机所接管的工作切换回已经修复的异常机。

2.2.2 双工方案(Duplex)

系统中两台主机皆为工作机，它们并行同步响应外部服务请求，并同步处理。在处理过程中，分不同阶段对处理中间结果进行比对并给出每个中间结果的评估，在最后输出时根据所有评估和当前的最终结果表决策略得出唯一的输出结果。该方案在系统层和应用层都实现了高可靠性，但实现技术较为复杂。

2.2.3 双机热备方案(Hot Standby)

一台主机为工作机，另一台主机为备份机，在系统正常工作情况下，工作机和备份机同时处理外部请求，但只有工作机响应请求，对外提供服务。工作机和备份机互相监听对方的健康信息。当工作机出现异常时，备份机主动接管工作机的工作，继续对外提供服务，从而保证系统能够不间断的运行。此时，原来的备份机就成了工作机，而原来的工作机就成了备份机。待工作机经过修复正常后，系统管理员通过管理命令或经由以人工或自动的方式将其切回系统，经过与工作机数据同步后，以备份机身份开始继续工作，同时开始与工作机开始心跳信息交互。

2.3 集群技术

计算机集群是一组物理上通过高速互连网络连接在一起的计算机集合，通过附加的集群系统软件，互相协作，作为一个整体对外提供服务。集群系统按照使用的不同要求主要分为高可用性集群和高性能集群。

高可用性集群的主要功能就是提供可靠的、不间断的服务。许多应用程序都需要二十四小时连续运行，如军事、电信、金融和网络等领域的应用等。对这些应用程序而言，暂时的停机都会导致数据的丢失和灾难性的后果。

集群系统具有良好的可用性。集群的某结点出现故障时，集群软件可以自动地将待处理的事务分配到正常的结点，继续向用户提供持续的服务。从用户的角度来开，集群系统具有单一映像，集群内的故障切换对用户而言是透明的。

集群系统还具有良好的扩展性。只需很少的配置工作就可以方便地向集群中加入或删除工作结点。

2.4 RAID技术

磁盘子系统是计算机系统的重要组成部分。磁盘故障将直接导致数据的丢失或破坏。有统计表明，磁盘故障占计算机系统故障的50%左右。因此，提高磁盘子系统的可靠性就可以大大提高计算机系统的可靠性。现在常采用廉价磁盘冗余阵列（RAID：Redundant Arrays of Inexpensive Disks）来实现高可用的磁盘系统。磁盘阵列把多个磁盘驱动器连接在一起协同工作，大大提高了速度，同时把磁盘系统的可靠性提高到接近无错的境界。

根据Patterson教授的定义，RAID可分为RAID0至RAID5共6个等级。其中除RAID0之外，都具有容错功能，单盘失效故障不会影响数据的完整性和正常访问。而RAID5结合了磁盘分段和奇偶冗余技术的优点并具有较高的磁盘利用率，成为首选的高可用磁盘系统解决方案。

3 程序设计中的高可用性技术

构建一个高可用的计算机系统不仅需要可靠的硬件支持，程序设计中对系统可用性的考虑和实现也非常重要，软硬结合才能真正实现系统的高可用性。在高可用系统中经常综合使用故障检测、故障诊断、故障恢复等技术手段保证系统的连续工作。

故障检测技术是提高计算机系统可用性的重要一步。故障检测应该具有及时准确的发现故障，占用系统开销小的特点。计算机系统的故障可分为：

3.1 硬件系统故障

如：系统总线错

I/O适配卡错；

CPU温度异常；

电源温度异常；

主板温度异常；

电源电压异常；

风扇电压异常；

风扇转速异常；

磁盘系统读写异常；

3.2 软件系统故障

如：系统内存使用超限；

CPU负载超限；

系统响应时间超限；

文件系统错；

应用程序逻辑错误；

根据故障出现的概率，又可将系统故障分为：

3.3 偶然故障

由偶然事件引发的程序执行过程中出现的难以重复出现的故障

3.4 永久故障

系统中重复出现的故障

故障检测技术就是通过对计算机系统的软硬件工作状态进行周期检测，建立系统健康报告。一旦发现故障，立即判定故障的性质。一般情况下，故障检测即可以判定故障的位置，如果必要可以使用故障诊断技术，准确判定故障的位置和性质。故障如果是永久故障，即故障重复出现，无法通过时间冗余的方法避免，则必须进行故障恢复操作，使系统在尽可能短的时间内重新恢复到正常工作状态，避免意外停机。一般硬件系统故障往往造成永久故障，因此在系统硬件设计时，对关键部分采用部件冗余设计，可以有效地避免此类故障。

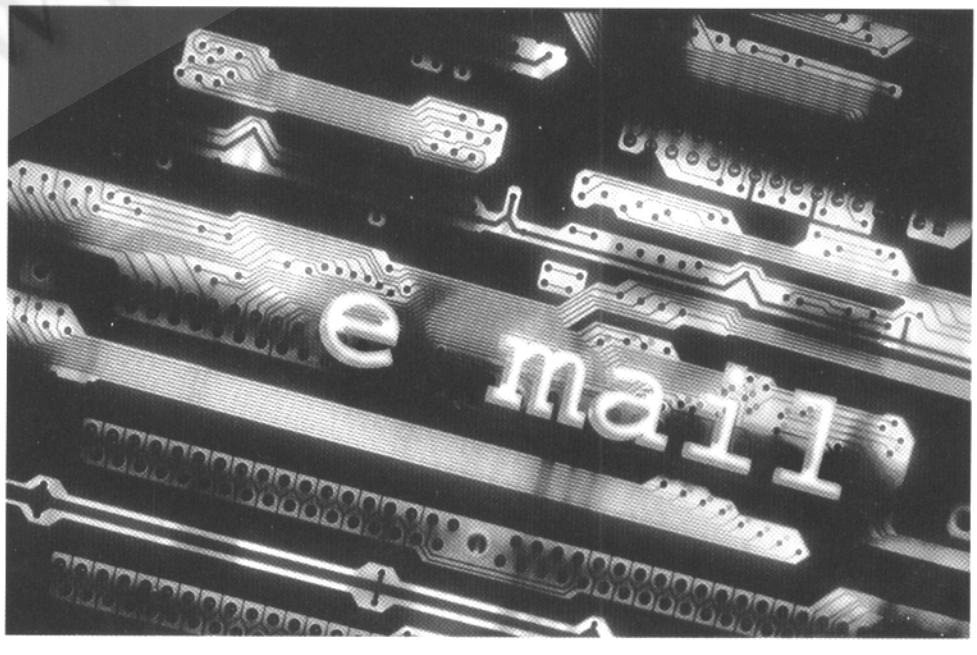
在实现中经常采用以下的检测方法：

3.5 心跳信息检测

心跳信息（Heartbeat）是一种多机系统间的监听机制。节点通过互连通道周期地对外广播本节点的健康状态（通常为“存活”状态），并接收其它节点的健康状态。互连通道可以采用RS232、Ethernet、FDDI或Memory channel等方式。心跳信息检测指的是周期地发送和接收本节点和其它节点的心跳信息，根据心跳信息的情况判断当前系统的工作状态，当发现心跳信息异常或心跳信息超时时，采取相应措施，保证系统能够正确连续地工作。心跳信息检测的检测周期是心跳信息检测的关键问题。设置的间隔时间过于频繁，将会影响到系统的正常运行，占用系统资源；而设置的间隔时间太长，则检测会比较迟钝，影响故障检测的及时性，最终影响到整个系统运行的正确性。

3.6 系统定时巡检

在计算机系统的正常运行过程中，运行故障检测程序，周期性地对系统软硬件工作状态进行巡检，得到系统的健康报告。巡检的结果可以作为心跳信息的主要内容传递给其他节点，也可作为本机工作状态转换的判定条件报告给系统管理员。系统巡检的关键问



题有两个：一是故障检测覆盖率，二是故障检测判定阈值的设定。根据文献[4]，故障检测覆盖率是影响计算机系统可靠度的重要因素，提高故障检测覆盖率可以提高计算机系统的可靠性，尽可能多的检测系统可能的故障点，对于及早发现故障，对提高系统可用性具有直接意义，但是，随着故障检测覆盖率的提高，检测的难度和开销也响应增大，因此应选取适当的故障检测覆盖率。对于检测出的异常情况，如何判定是否属于故障也是故障检测的一个难点。检测阈值过宽，会造成故障的漏判，从而影响系统的可用性。检测阈值过于严格，又有可能造成故障的虚判，影响到系统的正常运行，在多机系统中甚至可能造成频繁切换，系统开销巨增，无法正常运行。检测阈值通常可以结合具体应用的特点，采用动态设置法和静态设置法。

故障恢复技术是建立在故障检测基础上的处理手段。常用的故障恢复技术可以分为前向恢复和后向恢复。

3.7 前向恢复

前向恢复技术指的是系统从故障中恢复时，从出错时刻以后的某一时刻点开始恢复。采用前向恢复时，事务的执行时间要比后向恢复短，任务的完成时间可预测性强。但是，选择恢复的时机通常比较困难，状态恢复困难。

3.8 后向恢复

后向恢复技术指的是系统从故障恢复时，退回到以前的某一个状态，重新开始处理。后向恢复通过在系统正常运行过程中设置检查点，保存系统当时的一致性状态，并对各进程进行相关性跟踪和记录。系统发生故障后，将相关进程回卷到故障前系统一致性状态(检查点)，经过状态恢复后从该检查点处重新执行，实现对系统故障的恢复，节省了大量重复计算时间，提高了计算机系统的可用性。

4 应用实例

我们在工程实践中，采用双机热备方案并结合部件冗余、集群技术、RAID技术、故障检测和故障恢复技术实现了一个高可用计算机系统。系统构成见图1。

本系统采用双机热备的机制工作，一台Server为工作机，一台Server为备份机。正常工作时，两个Server上运行相同的任务，工作机对外提供服务；备份机不输出，通过MC检测对方机心跳信息。当工作机发生故障时，将工作机降为备份机或切出进行维修，同时备份机转变为工作机，继续对外提供服务。两台Server通过MC构成一个集群，通过集群软件自动负载均衡。系统配置一个IP地址，集群对用户提供单一映像。

两台Server通过SCSI线缆连接到磁盘阵列柜，共享磁盘阵列使磁盘子系统的可用性大大提高。系统中Switch、MC Hub和每台Server上的网卡都采用部件冗余，双套互为备份。

在每台Server上都运行自行开发的故障检测和故障恢复进程，周期性地检测本机的软硬件工作状态，按照一定的策略判定故障性质，填写心跳信息发送给对方机，同时监听对方的心跳信息。

当检测到本机发生故障，无法继续正常运行时，如果通过心跳信息得知对方机工作正常，则通过故障恢复进程切换工作机和备份机。双机的数据同步通过MC高速通道进行传递。在传递过程中使用检查点进行后向恢复，使系统的服务保持连续，用户完全感受不到主备机切换的影响。

5 结束语

高可用计算机系统在一些关键业务中的应用将会越来越广泛。本文研究了实现高可用计算机系统的一些技术方法，将这些方法综合运用到工程实际中，较好的满足了系统高可用、连续服务的需求，取得良好的效果。

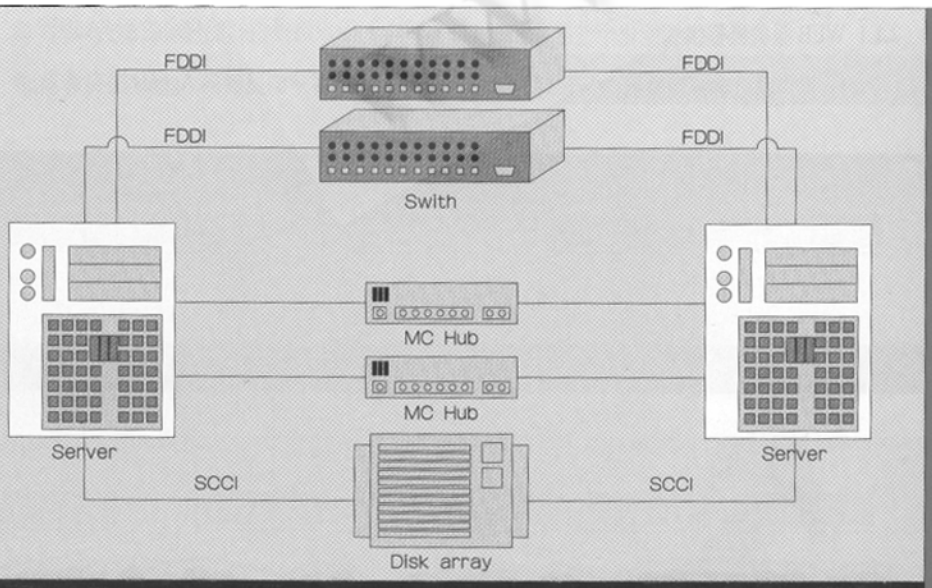


图1 高可用系统示意图

参考文献

1 顾宏等，高性能计算机系统的研究和发展，计算机工程与应用，1996.3:32。

2 蒋伟进等，网络可靠性体系结构研究与设计，计算机工程与应用，2000.12:156-169。

3 高文等，基于生灭过程的机群系统高可用性分析与设计。

4 高继祥等，双机热备计算机联锁系统可靠性与安全性指标分析，北方交大学报，1998.10.V22。